

Developing Dynamic Web-GIS
based Early Warning System for
the Communities at Landslide
Risks in Chittagong Metropolitan
Area, Bangladesh

Report
On
Land Cover Modelling



BUET-Japan Institute of Disaster Prevention and Urban Safety (BUET-JIDPUS)

Bangladesh University of Engineering and Technology (BUET), Dhaka-1000, Bangladesh

Email: jidpusinfo@jidpus.buet.ac.bd, Phone: +8802-9662975, Fax: +8802-9662975

Land Cover Modeling Report

Submitted to

International Centre for Integrated Mountain Development (ICIMOD)

Submitted by

BUET-Japan Institute of Disaster Prevention and Urban Safety
(BUET-JIDPUS); Bangladesh University of Engineering and
Technology (BUET), Dhaka-1000, Bangladesh

Author

Bayes Ahmed

Institute for Risk and Disaster Reduction
University College London (UCL), UK
Email: bayes.ahmed.13@ucl.ac.uk

January 2015



1. Background

Land use and land cover changes have been recognized as one of the most important factors stirring rainfall-triggered landslides [1]. Numerous researchers as a response to land use/ land cover changes have identified landslide incidence. Thomas Glade (2002) has showed how landslide can take place because of the change in land use in the context of New Zealand [1]. Mugagga *et al.* (2011) has also depicted the impacts of land use changes and its implications for the occurrence of landslides in mountainous areas of Eastern Uganda [2]. Moreover, expansion of urban development, deforestation and increased agricultural practices into the hillslope areas are immensely threatened by landslide hazards [3, 4, 5].

Land cover changes (e.g. deforestation) cause large variations in the hydro-morphological functioning of hillslopes, affecting rainfall partitioning, infiltration characteristics and runoff production. All these factors trigger landslides in hilly areas [6, 7]. Therefore it can be stated that there is a strong and positive correlation between land-use change and landslides. At this drawback, this report tries to project the future land cover change of Chittagong Metropolitan Area (CMA).

2. Study Area

The selected study area is a part of Chittagong Hill Tracts (CHT) of Bangladesh (Figure 1). It is important to know about the characteristics of CHT, to understand the causes and geological reasons of landslides in CMA. Therefore, the extent of the study area for land cover modelling is chosen as CHT. The total area of CHT is 19887.70895 square kilometres (sq. km.) [Reference system: UTM-46N].

3. Land Cover Mapping

Landsat Thematic Mapper (TM) satellite images are used for the land cover mapping of CHT area. Initially four scenes were collected to cover the whole CHT area. TM sensor collects reflected energy in three visible bands (blue = 1, green = 2, and red = 3) and three infrared bands (two NIR = 4, 5 and one middle infrared = 7). The base years for land cover mapping are selected as 1990, 2000 and 2010.

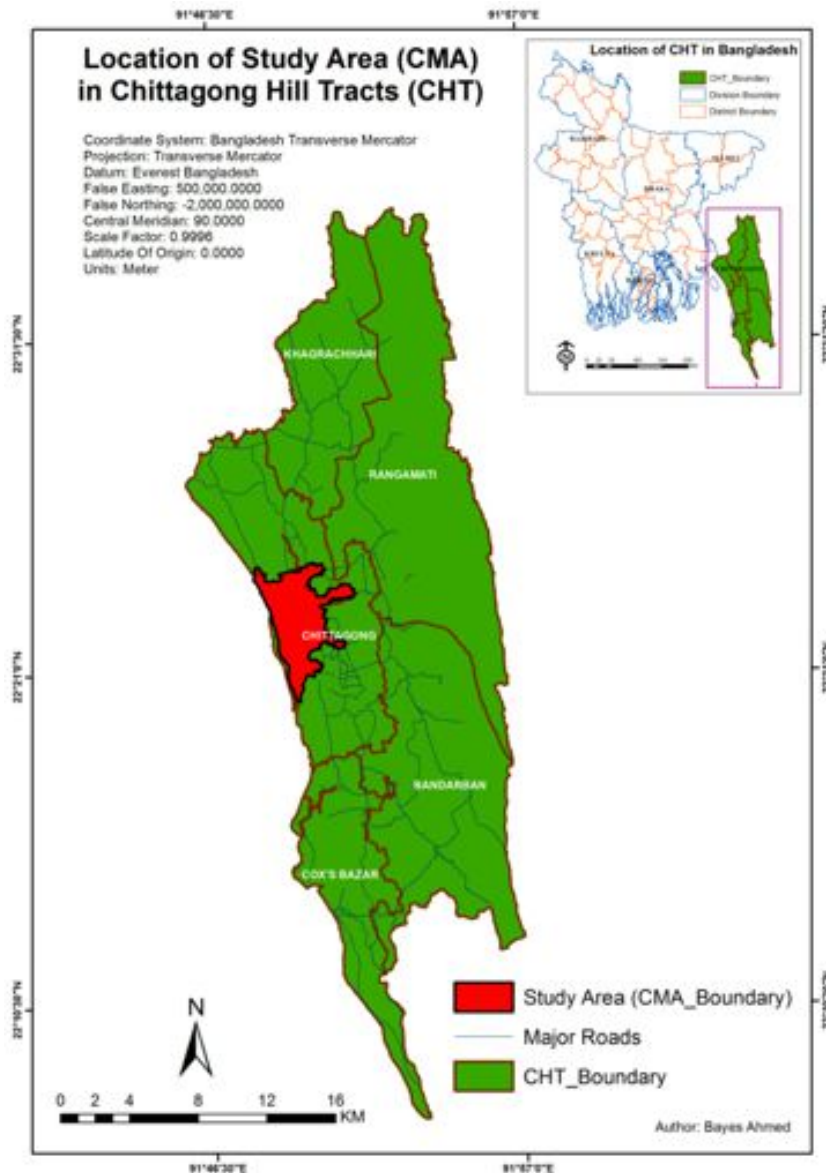


Figure 1. Location of Chittagong Metropolitan Area in Chittagong Hill Tracts

Among the four scenes, three were acquired using the Global Visualization Viewer (GLOVIS) of United States Geological Survey (USGS) and the one was from GISTDA (Geo-Informatics and Space Technology Development Agency), Thailand. However, thermal band was not used in this particular study. The details of the scenes used are listed in Table 1. All the image dates are of the dry season in Bangladesh.

The land cover classification methodology for this research is based on ‘Object Based Image Analysis (OBIA)’. ‘OBIA’ is also called ‘Geographic Object-Based Image Analysis (GEOBIA)’. ‘OBIA’ is a sub-discipline of geo-information science devoted

to partitioning remote sensing imagery into meaningful image objects, and assessing their characteristics through spatial, spectral and temporal scale. The fundamental step of any object based image analysis is a segmentation of a scene representing an image into image objects [7, 8].

Table 1. Details of the Landsat 4-5 TM scenes of CHT

Satellite	Sensor	Path	Row	Date (DD/MM/YY)	Source Agency
Landsat 4-5	TM	136	044	08/02/2010	USGS
		136	045	06/12/2009	
		135	045	01/02/2010	GISTDA
		135	046	01/02/2010	

The projection detail of all the raster images (cell size 30m × 30m)/ vector-shapefiles used in this report is as follows:

Projection: Bangladesh Transverse Mercator (BTM)

False Easting: 500000.000000

False Northing: -2000000.000000

Central Meridian: 90.000000

Scale Factor: 0.999600

Latitude of Origin: 0.000000

Linear Unit: Meter (1.000000)

Geographic Coordinate System: GCS_Everest_Bangladesh

Angular Unit: Degree (0.017453292519943299)

Prime Meridian: Greenwich (0.000000000000000000)

Datum: D_Everest_Bangladesh

Spheroid: Everest_Adjustment_1937

Semi-major Axis: 6377276.344999999700000000

Semi-minor Axis: 6356075.413140240100000000

Inverse Flattening: 300.801699999999980000

At first, the acquired Landsat TM images were inserted in ‘eCognition Developer 64 8.7’ software for processing. The “multi-resolution segmentation” algorithm was used which consecutively merges pixels or existing image objects that essentially identifies single image objects of one pixel in size and merges them with their neighbours, based on relative homogeneity criteria. Multi-resolution segmentations are groups of similar pixel values, which merge the homogeneous areas into larger objects and heterogeneous areas in smaller ones [8, 9].

During the classification process information on spectral values of image layers, vegetation indices like the Normalized Difference Vegetation Index (NDVI) and land water mask; which were created through band rationing, slope and texture information were used. Image indices are very important during the image classification. Image rationing is a “synthetic image layer” created from the existing bands of a multispectral image. This new layer often provides unique and valuable information not found in any other individual bands. Image index is a calculated result or generated product from satellite band/channels. It is help to identify different land cover from mathematical definition [8, 9].

NDVI is a standardized index that allows generating an image displaying greenness (relative biomass). NDVI was calculated using the formula: $NDVI = (NIR - red) / (NIR + red)$. This index output values between -1.0 and 1.0 , mostly representing greenness, where any negative values are mainly generated from clouds, water, and snow, and values near zero are mainly generated from rock and bare soil. Very low values of NDVI (0.1 and below) correspond to barren areas of rock or sand. Moderate values (0.2 to 0.3) represent shrub and grassland, while high values (0.6 to 0.8) indicate temperate and tropical rainforests [9].

Land and water mask: Land and water mask indices values can range from 0 to 255, but water values typically range between 0 and 50. The land and water mask was created using the formula: Land and water mask: $IR / (Green \times 100)$ [8].

The next step is to code these image objects according to their attributes, such as NDVI, land and water mask, layer value, colour and relative position to other objects using user-defined rules. In this process, selected object that represent patterns were

recognized with the help from other sources known as ‘ground-truth’ information and high-resolution Google earth images. Normally similar features observe similar spectral responses and are unique with respect to all other image objects [9].

After that comparison, features using the ‘2D Feature Space Plot’ were used for correlation of two features from the selected image objects. Developing rule sets investigated single image objects and generated land cover map. Image objects have spectral, shape, and hierarchical characteristics and these features are used as sources of information to define the inclusion-or-exclusion parameters used to classify image objects. Over each scene, rules were generated for each land cover class and evaluated for their separation, tested for their visual assessment over Google earth images [9].

After ascertaining the class separation using segment-based approach, classification is performed to get land cover classification map for each scene. Each scene thus prepared was evaluated again with available field data and Google earth image over randomly selected points for accuracy assessment. After finalization of classification of each scene, all the scenes were gone through mosaic to obtain land cover map of CHT area (Figures 2-5). For this research purpose, 5 broad land-cover classes (urban area, semi-urban area, water body, vegetation and bare soil) were chosen by reclassification technique. Here the land coverage is as follows:

Builtup Area = Urban and rural area, settlements and transportation infrastructure
Vegetation = Shrub land, rubber plantation, rained herbaceous crops, mangrove plantation, irrigated herbaceous crops, and crop in sloping land/tea
Water Body = River, reservoir/ponds, lake, canal, Bay of Bengal and low land
Hilly Forest = Hilly areas, forest in hills and highly dense rain forest
Bare Soil = Sand, sea beach, fallow land, earth and sand land in-fillings, open space, bare and exposed soils etc.

The overall accuracies of the lands cover images (1990, 2000 and 2010) were found approx. 89%, 92% and 88.6% respectively.

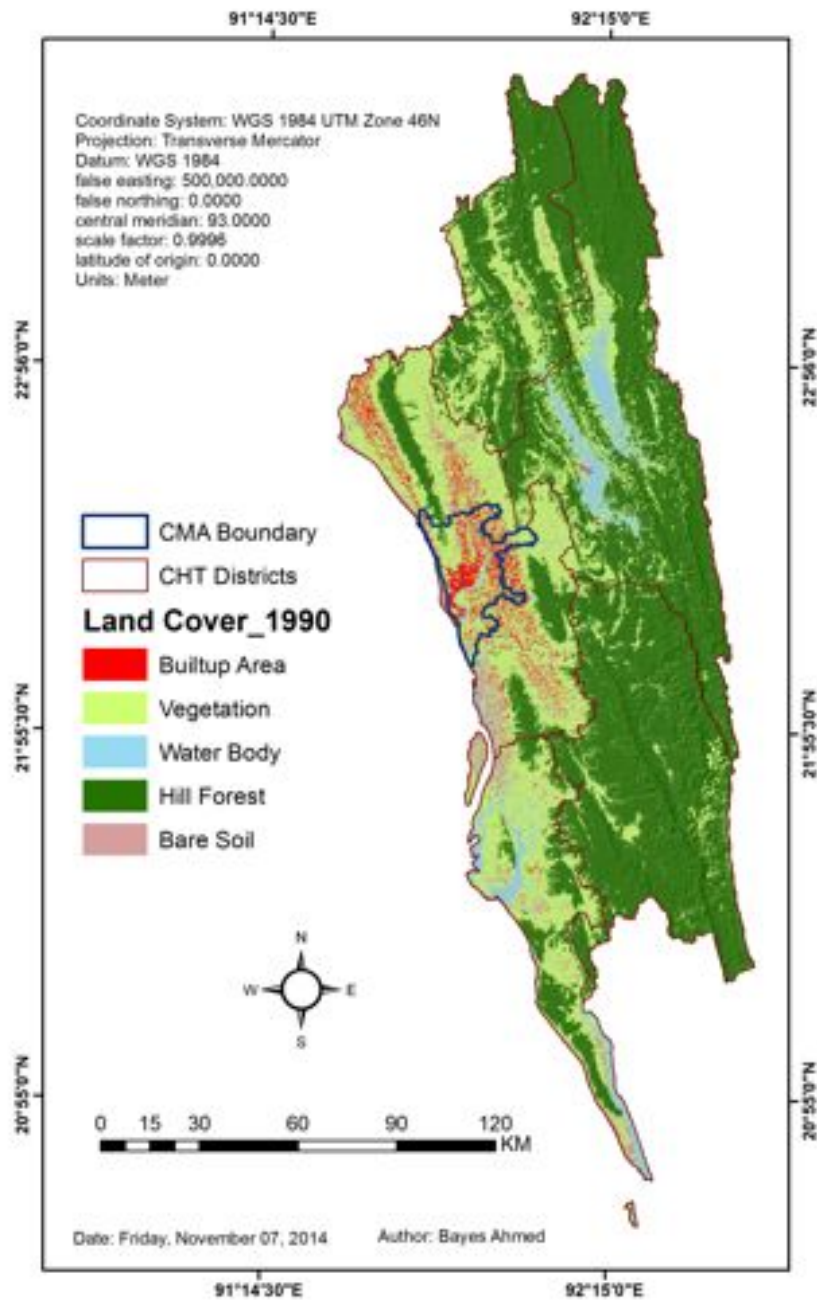


Figure 2. Land Cover Map of CHT (1990)

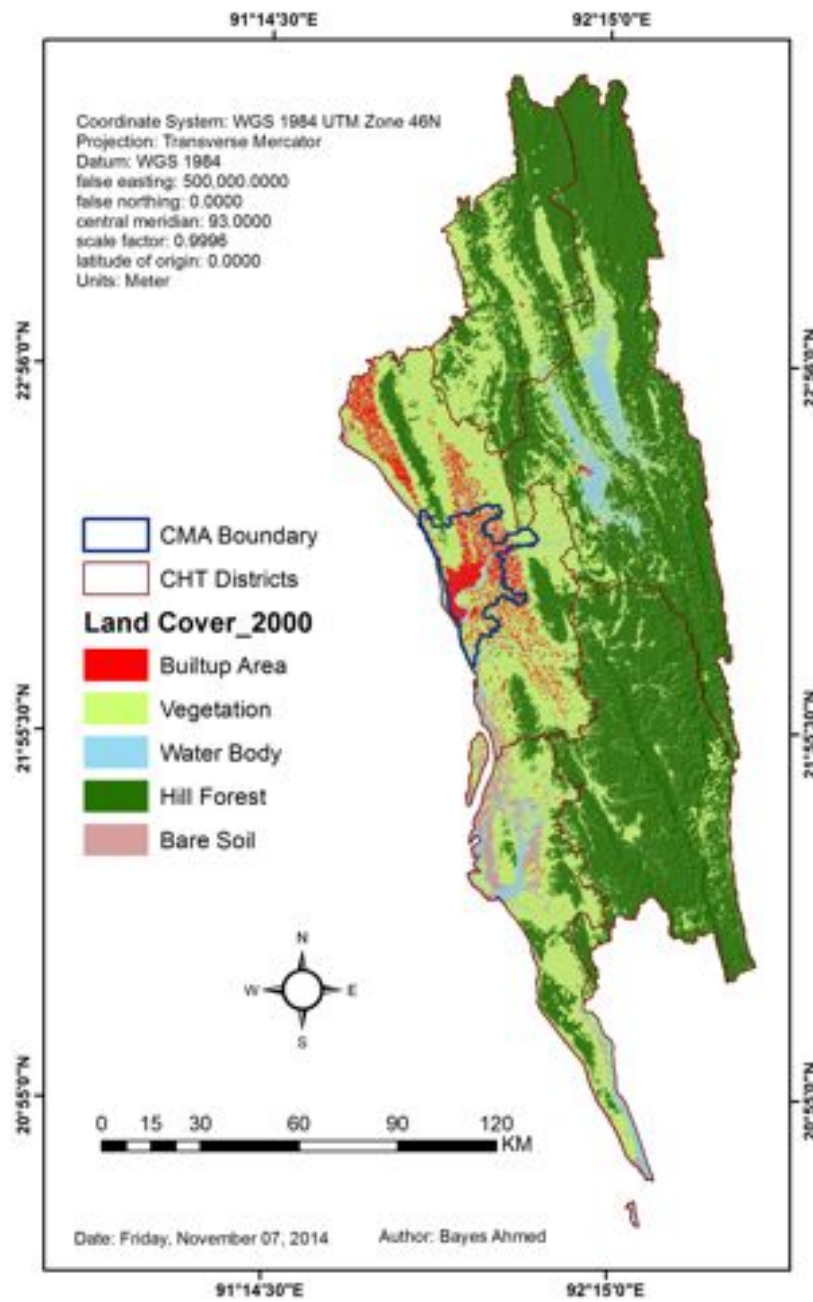


Figure 3. Land Cover Map of CHT (2000)

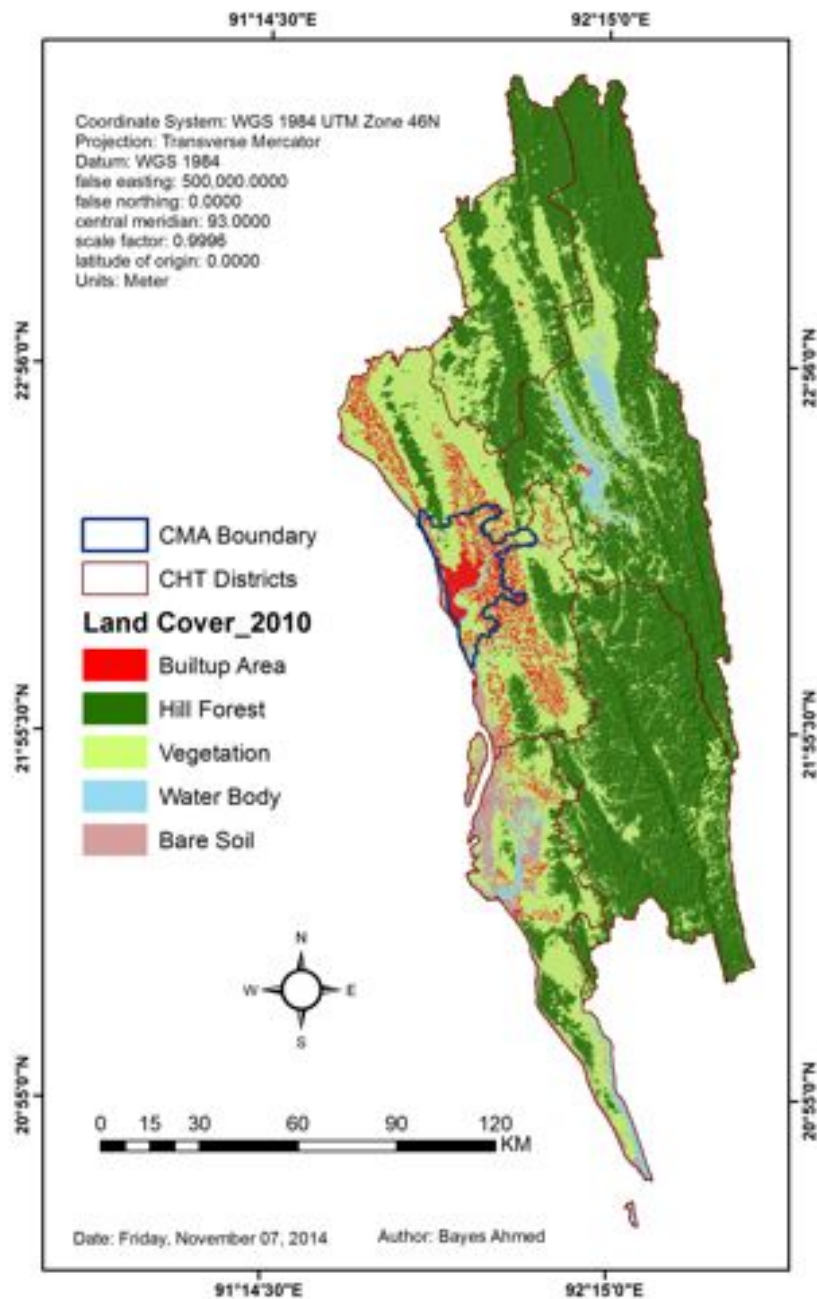


Figure 4. Land Cover Map of CHT (2010)

4. Change Detection

In remote sensing, “Change Detection” is defined as the process of determining and monitoring the changes in the land cover types in different time periods. It provides the quantitative analysis of the spatial distribution in the area of interest [10]. The percentage of change in area is depicted in Table 2.

Table 2. Percentages of Presence of Land Cover Types in CHT (1990-2010)

Land Cover	1990	2000	2010
	% Of Area		
Hill Forest	58	55	54
Vegetation	32	35	35
Builtup Area	3	3	4
Water Body	4	5	3
Bare Soil	3	2	3

Here the change detection is performed between 1990 and 2010. Figure 5 is showing the major transitions (by ignoring transitions less than 0.5 sq. km.) among different land cover types. The figures showing the changes (gains, losses and persistence) for each land cover type are attached in Appendix-I. Some major changes are as follows:

- i. 1128.2664218 sq. km. areas converted from Hill Forest to Vegetation
- ii. 421.1719313 sq. km. areas converted from Vegetation to Builtup Area
- iii. 281.4113228 sq.km areas were converted from Vegetation to Bare Soil
- iv. 108.1640835 sq.km were converted from Water Body to Bare Soil

The following changes are also notable:

- i. Gains in vegetation (approx. 700 sq. km.) and builtup area (approx. 380 sq.km.) were highest (Figures 6 and 7)
- ii. Hill forest is decreasing alarmingly (780 sq. km) and the major contributor to it is the vegetation land cover type (Appendix-II).
- iii. Bare soil (100 sq. km) and water bodies (80 sq. km) were converted to vegetation type (Appendix-II).
- iv. Vegetation (320 sq. km) and bare soil (35 sq. km) were the major contributors to convert into builtup area (Appendix-II).
- v. Approximately 115 sq. km and 80 sq. km water body type were converted to vegetation and bare soil respectively (Appendix-II).
- vi. Bare soil was converted to vegetation (140 sq. km) and builtup area (30 sq. km).

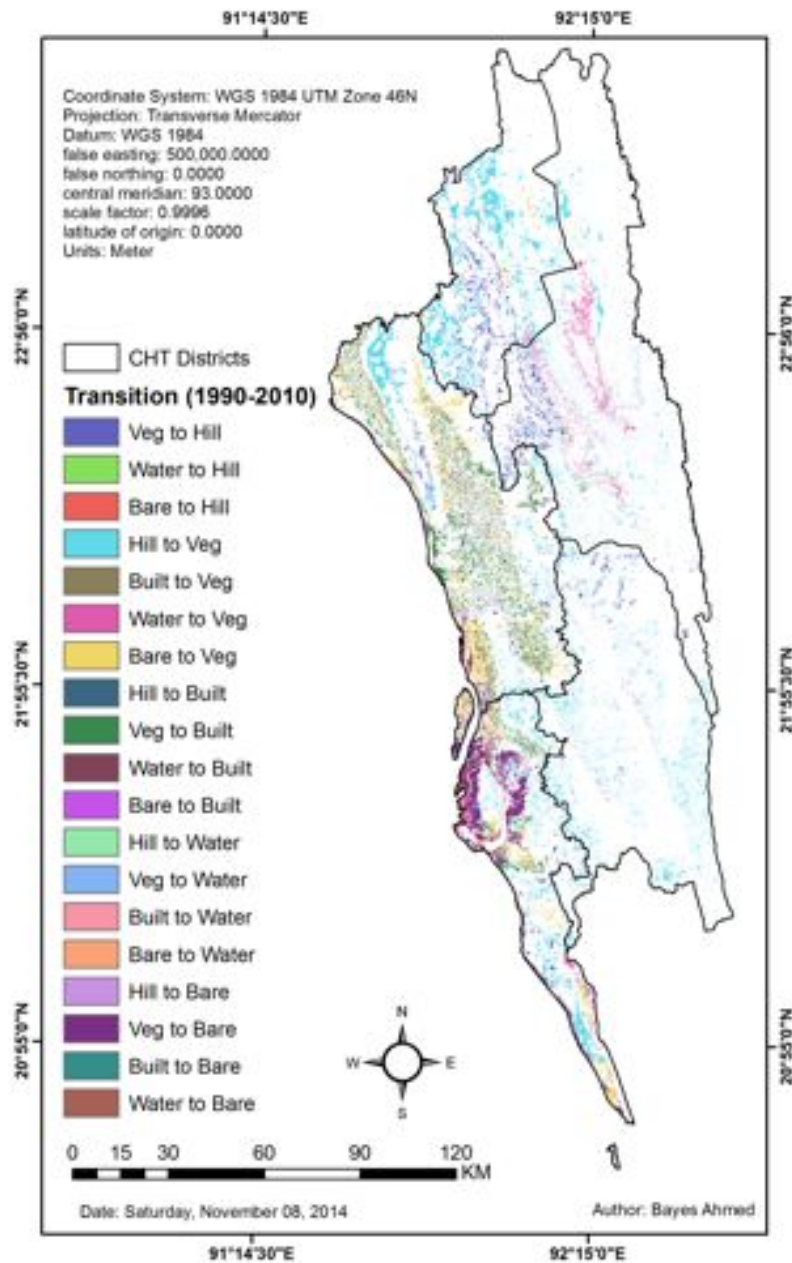


Figure 5. Transitions among Different Land Cover Types (1990-2010)



Figure 6. Gains and Losses between 1990-2010

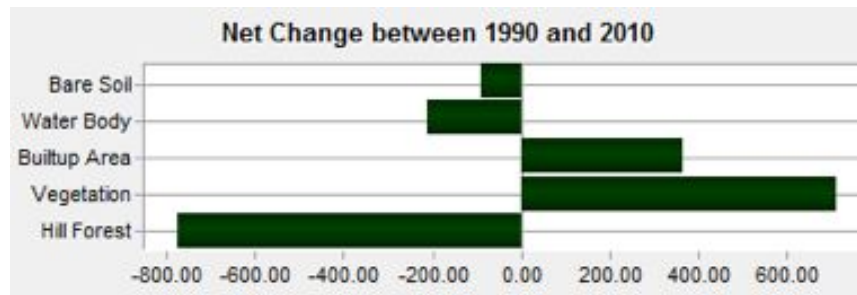


Figure 7. Net Change between 1990-2010

5. Land Cover Modelling Using Multi Layer Perceptron Markov Model

Multi Layer Perceptron (MLP) Markov Model is used to project the land cover map of CHT for 2030. This model combines both the Markov Chain and Artificial Neural Network (ANN) techniques [11, 12].

5.1. Markov Chain

A Markov chain is a stochastic process (based on probabilities) with discrete state space and discrete or continuous parameter space [13]. In this random process, the state of a system *s* at time (*t+1*) depends only on the state of the system at time *t*, not on the previous states.

5.1.1. Markov Property

In a Markov chain the probability of the next state is only dependent upon the current state. This is called Markov property and stated as [14]:

$$P(\xi_{t+1} = X_{i_{t+1}} | \xi_1 = X_{i_1}, \dots, \xi_t = X_{i_t}) = P(\xi_{t+1} = X_{i_{t+1}} | \xi_t = X_{i_t})$$

The probability of a Markov chain ξ_1, ξ_2, \dots can be calculated as [14]:

$$P(\xi_1 = X_{i_1}, \dots, \xi_t = X_{i_t}) = P(\xi_1 = X_{i_1}) * P(\xi_2 = X_{i_2} | \xi_1 = X_{i_1}) * P(\xi_t = X_{i_t} | \xi_{t-1} = X_{i_{t-1}})$$

The conditional probabilities:

$$P(\xi_{t+1} = X_{i_{t+1}} | \xi_t = X_{i_t})$$

These are called the “Transition Probabilities” of the Markov chain [14].

5.1.2. Transition Matrix for a Markov Chain

Let's consider a Markov chain with n states s_1, s_2, \dots, s_n . Let p_{ij} denote the transition probability from state s_i to state s_j , *i.e.*,

$$P(\xi_{t+1} = s_j | \xi_t = s_i)$$

The transition matrix ($n \times n$) of this Markov process is then defined as [14]:

$$P = \begin{bmatrix} p_{11} & \dots & p_{1n} \\ \dots & \dots & \dots \\ p_{n1} & \dots & p_{nn} \end{bmatrix}, p_{ij} \geq 0, \sum_{j=1}^n p_{ij} = 1, i = 1, \dots, n$$

Predictions of the future state probabilities can be calculated by solving the matrix equation [19]:

$$p(t) = p(t - 1) \cdot P$$

With increasing time steps, a Markov chain may approach to a constant state probability vector, which is called limiting distribution [19]:

$$p(\infty) = \lim_{t \rightarrow \infty} p(t) = \lim_{t \rightarrow \infty} p(0) \cdot P^t$$

5.2. Multi Layer Perceptron Markov Model

The term 'Artificial Neural Network (ANN)' has been inspired by human biological nervous system [15]. In a typical ANN model, simple nodes are connected together to form a network of nodes. Some of these nodes are called input nodes; some are output nodes and in between there are hidden nodes [16]. Multi Layer Perceptron (MLP) is a feed-forward Neural Network with one or more layers between input and output layers. The great advantage of using MLP perceptron neural network is that it gives the opportunity to model several or even all the transitions at once [12].

5.2.1. The Feed-Forward Concept of Multi Layer Perceptron Neural Network

MLP neural network uses the back propagation (BP) algorithm. The calculation is based on information from training sites [12]. Back propagation involves two major steps, forward and backward propagation. The input that a single node receives is weighted as:

$$net_j = \sum \omega_{ji} O_i$$

where, w_{ij} = the weights between node i and node j ; O_i = the output from the node i

The output from a given node j is computed as [16]:

$$O_i = f(\text{net}_j)$$

f = a non-linear sigmoid function that is applied to the weighted sum of inputs before the signal passes to the next layer

This is known as “Forward Propagation”. Once it is finished, the activities of the output nodes are compared with their expected activities. In normal circumstances, the network output differs from the desired output (a set of training data, e.g., known classes). The difference is termed as the error in the network [16]. The error is then back-propagated through the network. Now the weights of the connections are corrected as follows [12]:

$$\Delta\omega_{ji}(n+1) = \eta(\delta_j O_i) + \alpha\Delta\omega_{ji}(n)$$

η = the learning rate parameter; δ_j = an index of the rate of change of the error; α = the momentum parameter.

The process of the forward and backward propagation is repeated iteratively, until the errors of the network minimized or reaches an acceptable magnitude [16]. The purpose of training the network is to get proper weights both for the connection between the input and hidden layer, and between the hidden and the output layer for the classification of unknown pixels [12]. Several factors affect the capabilities of the neural network to generalize [16]. These include:

5.2.2. Number of Nodes

In general, the larger the number of nodes in the hidden layer, the better the neural network represents the training data [16]. The number of hidden layer nodes is estimated by the following equation [12]:

$$N_h = INT(\sqrt{N_i * N_o})$$

where, N_h = the number of hidden nodes; N_i = the number of input nodes; N_o = the number of output nodes

5.2.3. Number of Training Samples and Iterations

The number of training sample also affects the training accuracy. Too few samples may not represent the pattern of each category while too many samples may cause

overlap. Again too many iterations can cause over training that may cause poor generalization of the network [12]. Over training can be prevented by early stopping of training [15]. The acceptable error rate is evaluated based on the Root Mean Square (RMS) Error [15]:

$$RMS = \frac{\sum(e_i)^2}{N} = \frac{\sum(t_i - a_i)^2}{N}$$

where, N = the number of elements; i = the index for elements; e_i = the error of the i^{th} element; t_i = the target value (measured) for i^{th} element; a_i = the calculated value for the i^{th} element.

5.2.4. Multi Layer Perceptron Markov Modeling

The basic concept of modeling with MLP neural network adopted in this research is the following two criteria (Appendix-III):

- i. Changes from all land cover type to builtup area
- ii. Changes from hill forest to all other land cover types

These two criteria have been selected based on the trends of change detection over the years and also giving priority on the issues related to landslide vulnerability. Therefore, the following 9 transitions were selected for the MLP_Markov modeling:

- i. Hill Forest to Vegetation
- ii. Hill Forest to Builtup Area
- iii. Hill Forest to Water Body
- iv. Hill Forest to Bare Soil
- v. Vegetation to Builtup Area
- vi. Vegetation to Bare Soil
- vii. Water Body to Builtup Area
- viii. Water Body to Bare Soil
- ix. Bare Soil to Builtup Area

Next, three driver variables along with the Digital Elevation Model (DEM) are selected for MLP_Markov modeling (Appendix-IV). The driving variables are (1990–2010): Distance from all land cover type to builtup area, Distance from Hill Forest to all other land cover types, and the Empirical likelihood image.

The empirical likelihood transformation is an effective means of incorporating categorical variables into the analysis (Appendix-IV). It was produced by determining the relative frequency of different land cover types occurred within areas of transition (2000 to 2010). The numbers indicate the likelihood of changing into builtup area. The higher the value, the likeliness of the pixel to change into the builtup cover type is more [11].

Now it is important to test the potential explanatory power of each variable. The quantitative measures of the variables have been tested through Cramer's V [17]. It is suggested that the variables that have a Cramer's V of about 0.15 or higher are useful while those with values of 0.4 or higher are good [12]. In most of the cases, the Cramer's V were found statistically significant (Table 3).

Table 3. Cramer's V for the Driving Variable for Land Cover Modelling

Cover Class	Cramer's V		
	Distance (All to Built)	Distance (Hill to All)	Empirical Likelihood Image
Overall V	0.3736	0.2268	0.3936
Hill Forest	0.5706	0.3466	0.7250
Builtup Area	0.5057	0.2365	0.6975
Vegetation	0.3658	0.2133	0.2345
Water Body	0.2033	0.1390	0.0808
Bare Soil	0.1993	0.0590	0.0752

After getting satisfactory Cramer's V values for all the driving variables, now the turn is to run MLP neural network model. For this purpose, 10,000 iterations were chosen. The minimum number of cells that transitioned from 1990 to 2010 is 1589. Therefore, the maximum sample size has been chosen as 1,589 (50% training/ 50% testing). For each principal transition particular weights has to be obtained. The RMS error curve has been found smooth and descent after running MLP neural network (Figure 8). After all these combinations, the MLP running statistics gives a very high accuracy rate of 89.60%.

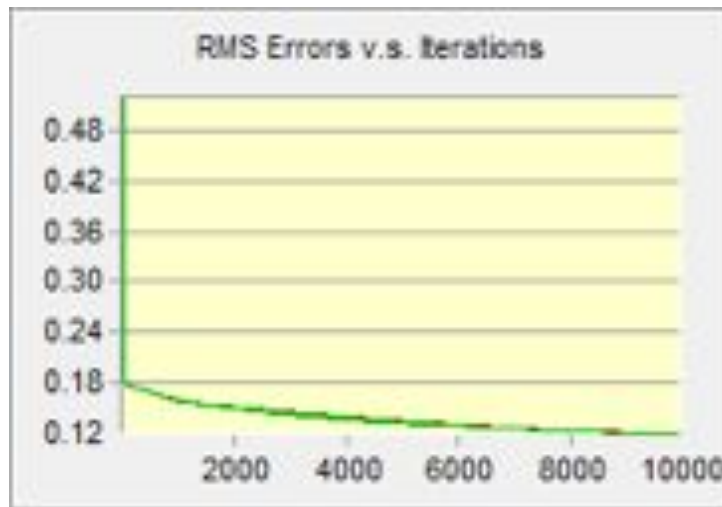


Figure 8. RMS Error Curve for MLP_Markov Modelling

Based on these running statistics the transition potential maps are produced (Appendix-V). These maps depict, for each location, the potential it has for each of the modelled transitions [12]. These are not the probability maps where the sum of values for a particular pixel location will not be 1. The reason behind this is because the MLP neural network outputs are obtained by applying fuzzy set to the signals into values from 0 to 1 with activation function (sigmoid). Here the higher values represent a higher degree of membership for that corresponding land cover type [12].

5.2.5. Future Prediction Using Multi Layer Perceptron Markov Model

Using this kind of MLP neural network analysis, it is possible to determine the weights of the transitions that will be included in the matrix of probabilities of Markov Chain for future prediction [11]. The transition probabilities are shown in Table 4. Based on all these information from MLP neural network, the final land cover map of 2030 (Figure 9) is simulated through Markov chain analysis. The projected areas of the land cover types in 2030 will be as follows (Figure 9):

Hill Forest = Approx. 9753 sq. km

Vegetation = Approx. 7300 sq. km

Builtup Area = Approx. 1362 sq. km

Water Body = Approx. 572 sq. km

Bare Soil = Approx. 893 sq. km

Table 5. Transition Probabilities Grid for Markov Chain for MLP_Markov Modelling

	Hill Forest	Vegetation	Builtup Area	Water Body	Bare Soil
Hill Forest	0.9020	0.0974	0.0001	0.0002	0.0002
Vegetation	0.0552	0.8204	0.0667	0.0133	0.0445
Builtup Area	0.0001	0.1892	0.8070	0.0020	0.0018
Water Body	0.0140	0.2345	0.0031	0.6233	0.1251
Bare Soil	0.0027	0.6816	0.0554	0.0447	0.2155

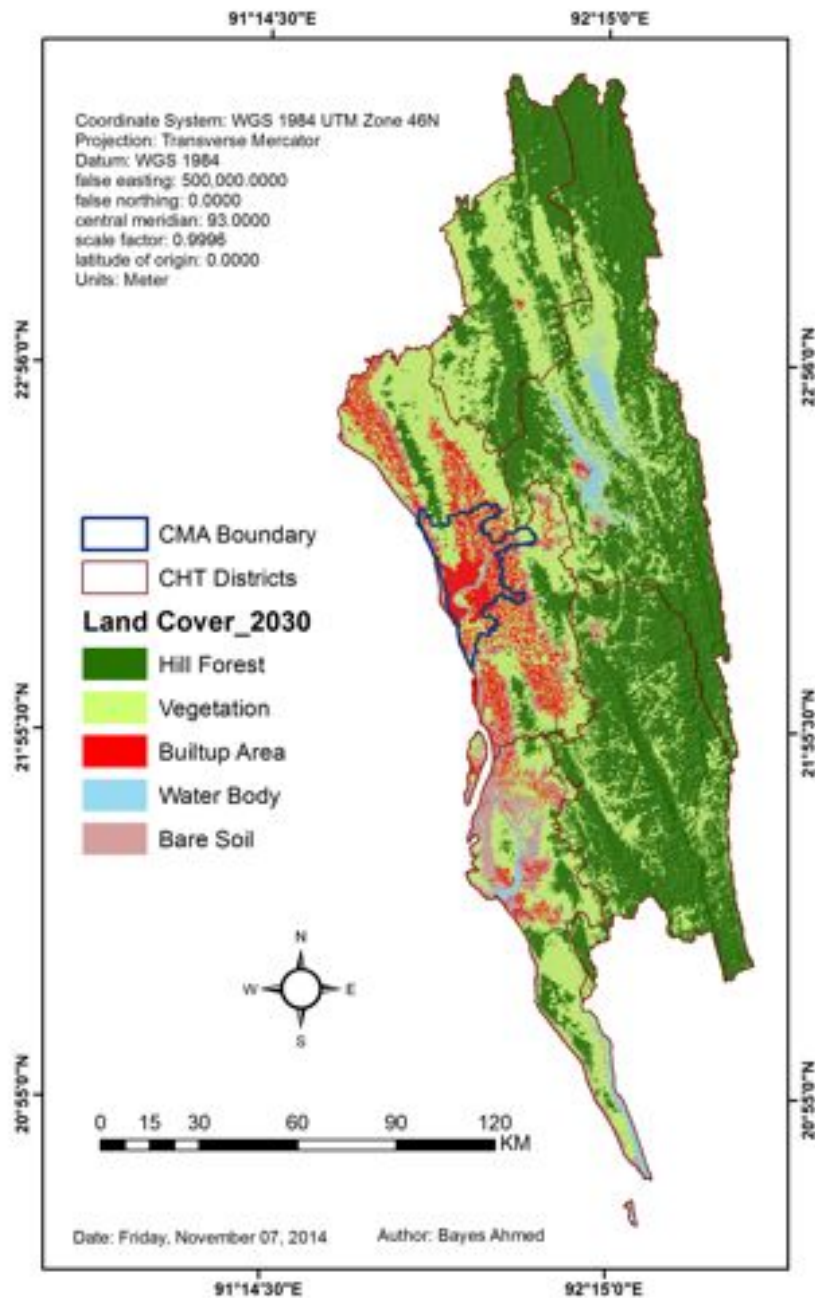


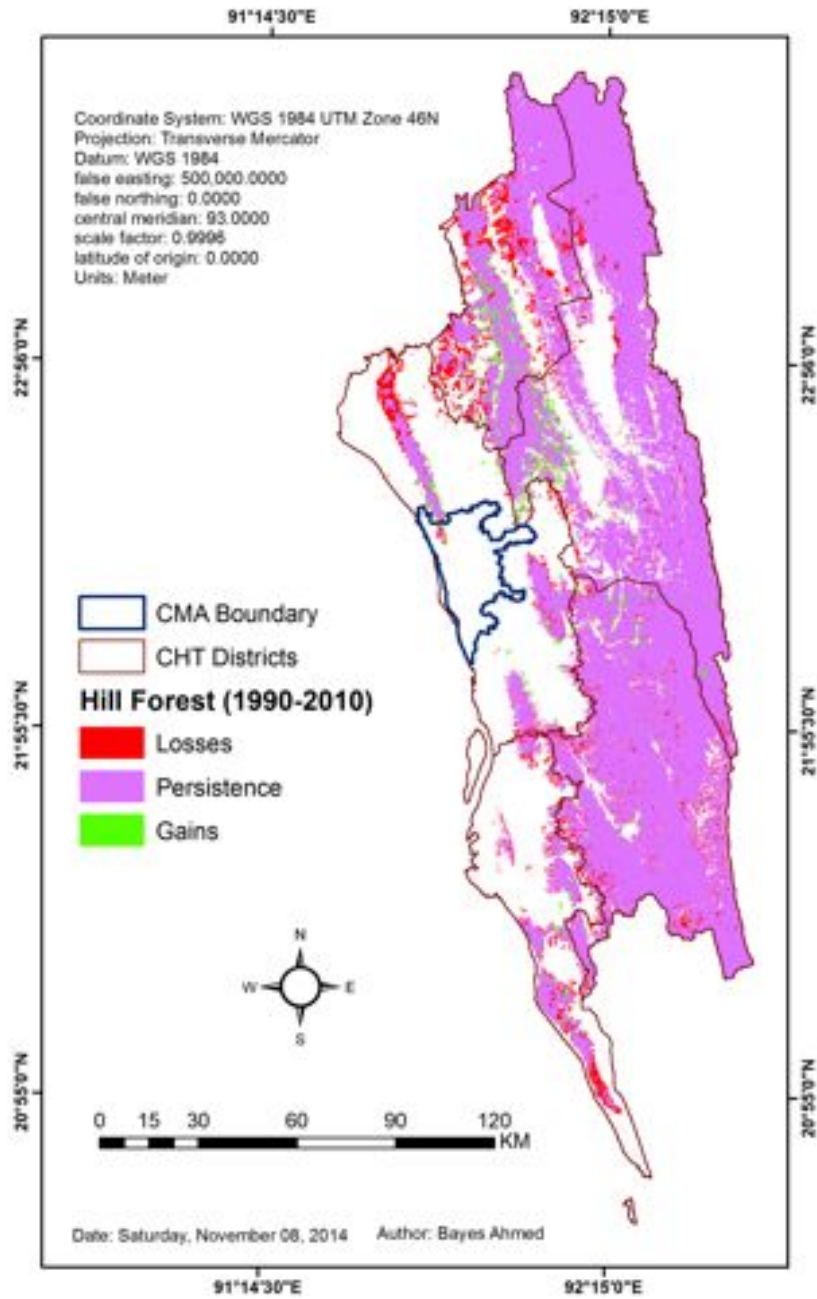
Figure 9. MLP_Markov Projected Land Cover Map of CHT (2030)

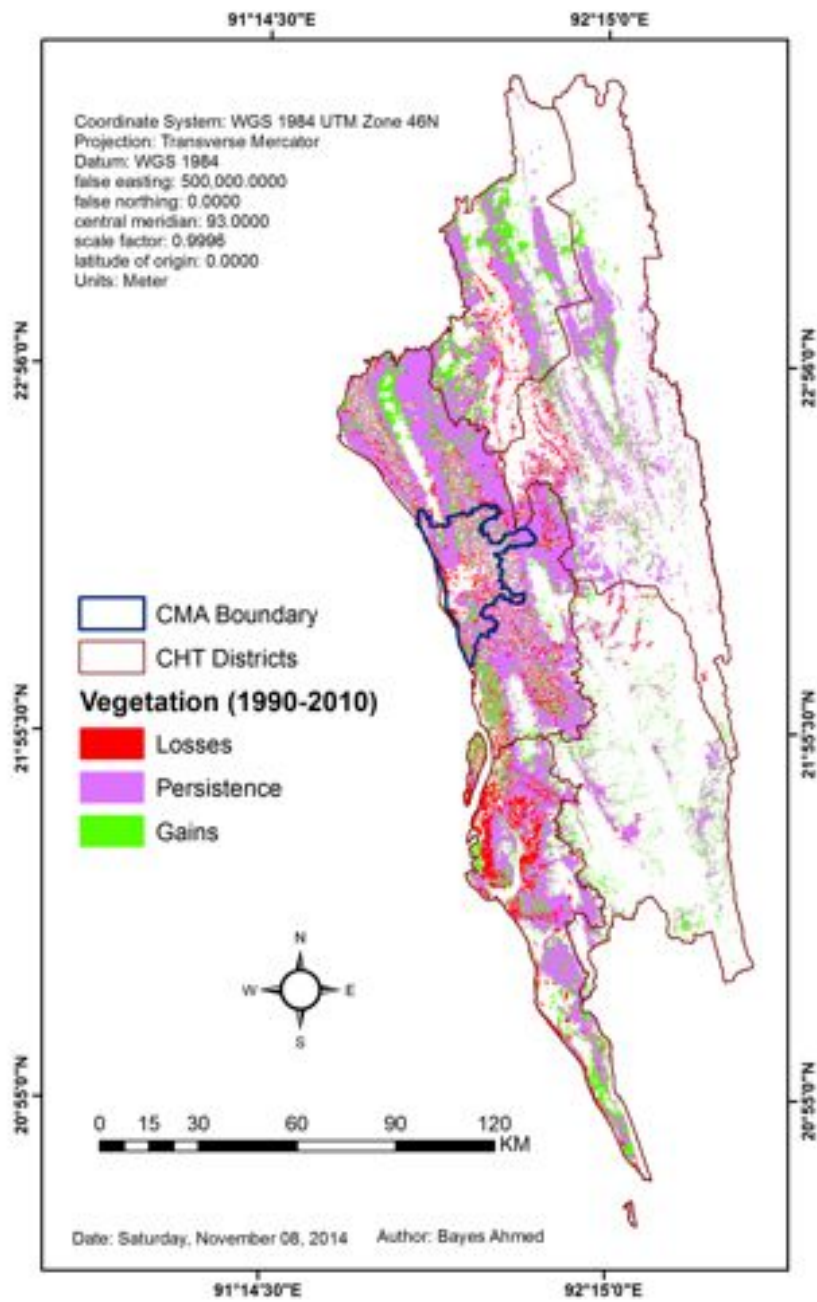
REFERENCES

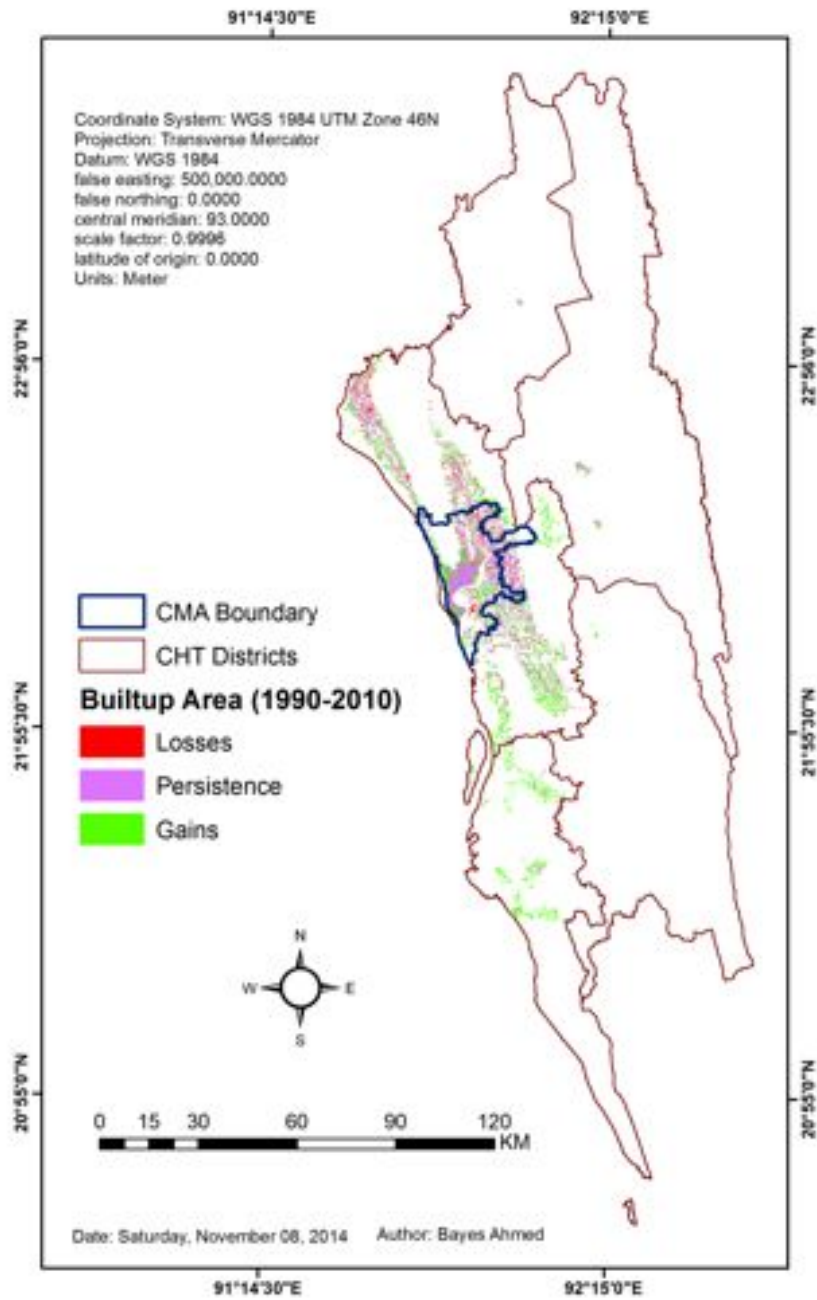
- [1] Thomas Glade, Landslide occurrence as a response to land use change: a review of evidence from New Zealand, *Catena*, 51 (2003) 297–314.
- [2] F. Mugagga, V. Kakembo and M. Buyinza, Land use changes on the slopes of Mount Elgon and the implications for the occurrence of landslides, *Catena*, 90 (2012) 39–46.
- [3] Xiaojun Yang and Liding Chen, Using multi-temporal remote sensor imagery to detect earthquake-triggered landslides, *International Journal of Applied Earth Observation and Geoinformation*, 12 (2010) 487–495.
- [4] I. Alcantara-Ayala, O. Esteban-Chavez and J.F. Parrot; Landsliding related to land-cover change: A diachronic analysis of hillslope instability distribution in the Sierra Norte, Puebla, Mexico; *Catena*, 65 (2006) 152–165.
- [5] H.L. Perotto-Baldiviezo, T.L. Thurow, C.T. Smith, R.F. Fisher & X.B. Wu; GIS-based spatial analysis and modeling for landslide hazard assessment in steeplands, southern Honduras; *Agriculture, Ecosystems and Environment*, 103 (2004) 165–176.
- [6] José M. García-Ruiz, Santiago Beguería, Luis Carlos Alatorre & Juan Puigdefábregas; Land cover changes and shallow landsliding in the flysch sector of the Spanish Pyrenees; *Geomorphology*, 124 (2010) 250–259.
- [7] Ahmed, B. (2014). Landslide susceptibility modelling using multi-criteria evaluation techniques in Chittagong Metropolitan Area, Bangladesh. *Landslides*. DOI 10.1007/s10346-014-0521-x
- [8] Uddin K. (2013). Image classification: hands on exercise using eCognition. eCognition Community, Trimble Geospatial Imaging, Arnulfstrasse 126, 80636 Munich, Germany.
- [9] Ahmed B, Rubel YA. (2013). Understanding the issues involved in urban landslide vulnerability in Chittagong metropolitan area, Bangladesh. The Association of American Geographers (AAG). Retrieved on: <https://sites.google.com/a/aag.org/mycoe-servirglobal/final-arafat>, assessed on 12 October 2014.
- [10] Tewolde, M.G.; Cabral, P. Urban sprawl analysis and modeling in Asmara, Eritrea. *Remote Sens.* 2011, 3, 2148-2165.

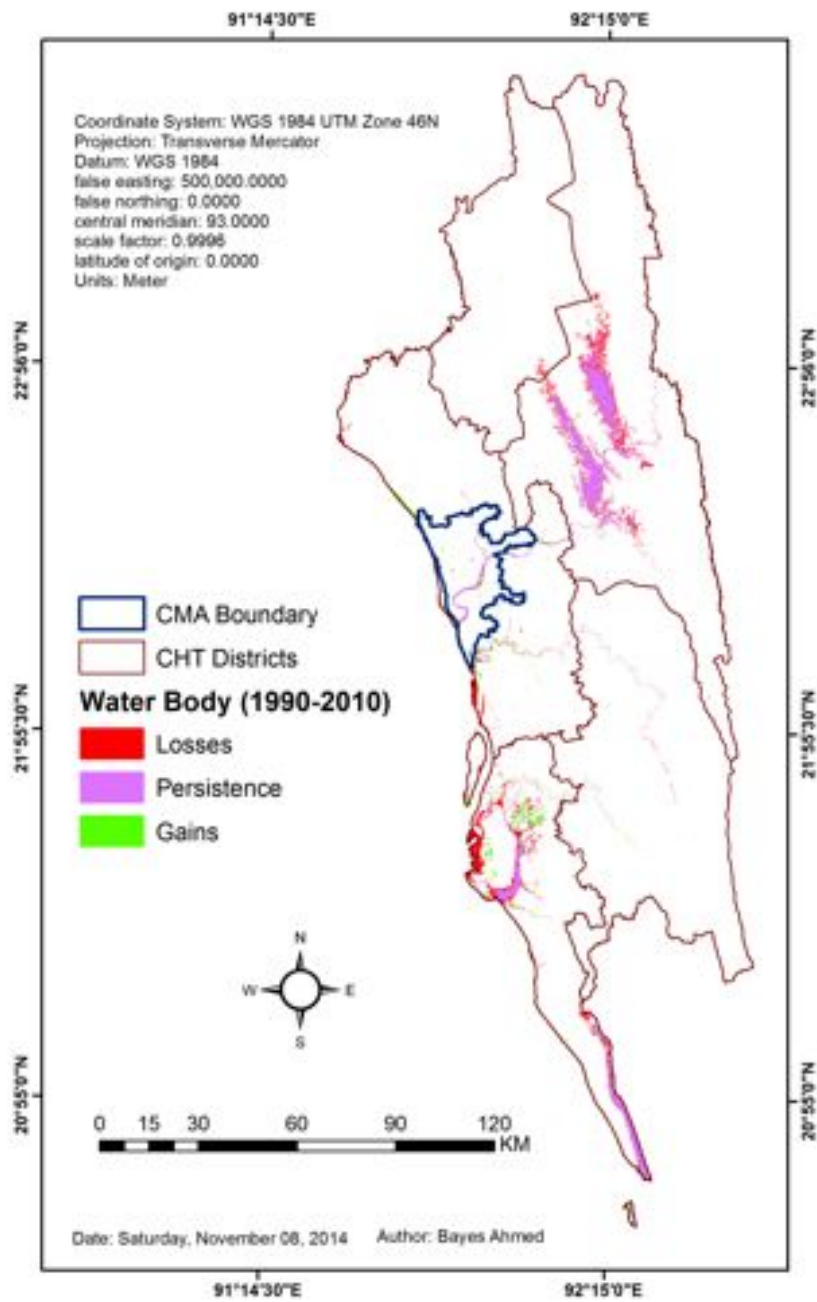
- [11] Ahmed B, Ahmed R (2012) Modeling urban land cover growth dynamics using multi-temporal satellite images: a case study of Dhaka, Bangladesh. *ISPRS Int J Geo-Inf* 1:3–31.
- [12] Eastman R (2012) *The IDRISI Selva Help*. Clark Labs, Clark University 950 Main Street, Worcester MA 01610–1477 USA.
- [13] Balzter, H. Markov chain models for vegetation dynamics. *Ecol. Model.* 2000, 126, 139-154.
- [14] Weng, Q. Land use change analysis in the Zhujiang Delta of China using satellite remote sensing, GIS and stochastic modelling. *J. Environ. Manage.* 2002, 64, 273-284.
- [15] Karul, C.; Soyupak, S. A comparison between neural network based and multiple regression models for chlorophyll-a estimation. In *Ecological Informatics*; Recknagel, F., Ed.; Springer: Berlin, Germany, 2006; pp. 309-323.
- [16] Atkinson, P.M.; Tatnall, A.R.L. Introduction neural networks in remote sensing. *Int. J. Remote Sens.* 1997, 18, 699-709.
- [17] Cramér, H. Methods of estimation. In *Mathematical Methods of Statistics*, 19th ed.; Chapter 33; Princeton University Press: Princeton, NJ, USA, 1999; pp. 497-506.

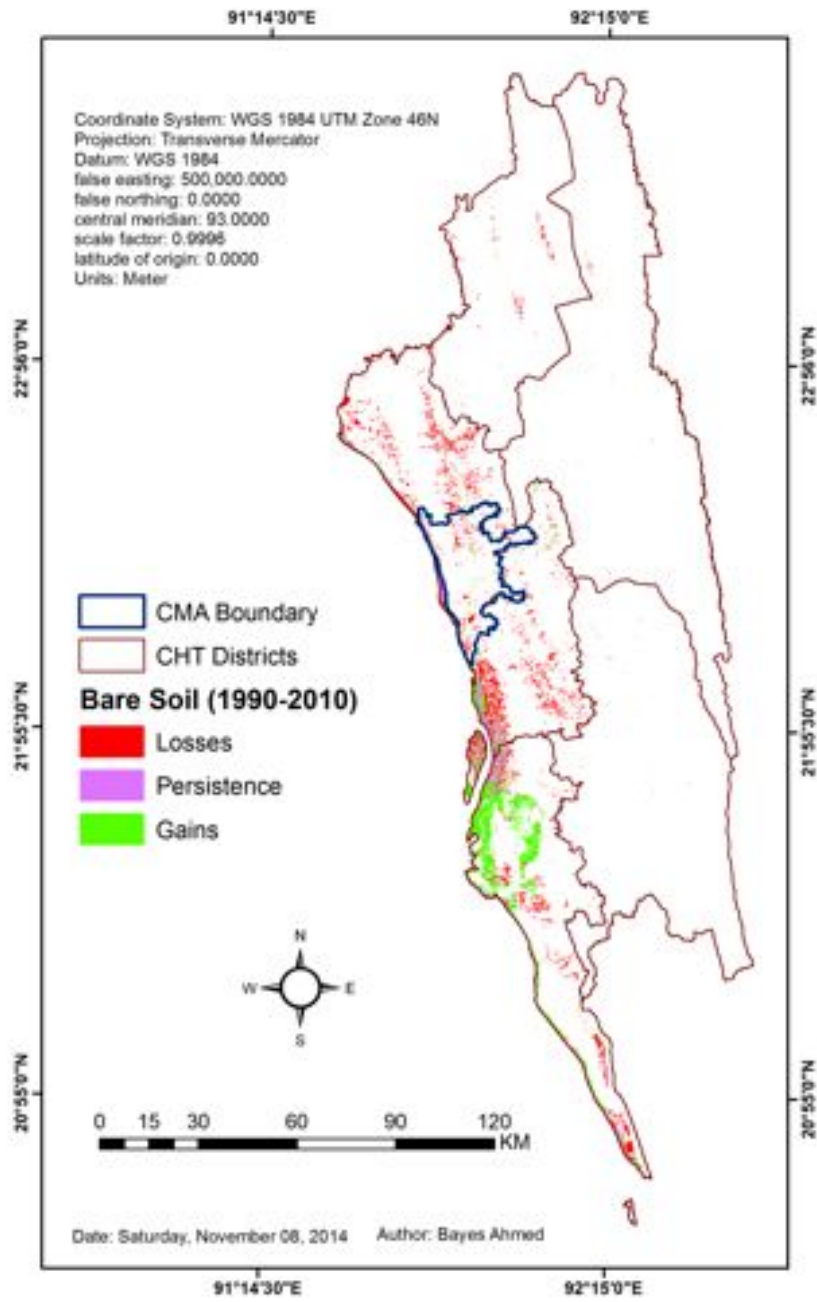
Appendix-I: Maps Showing the Transitions of Land Cover Types (1990-2010)



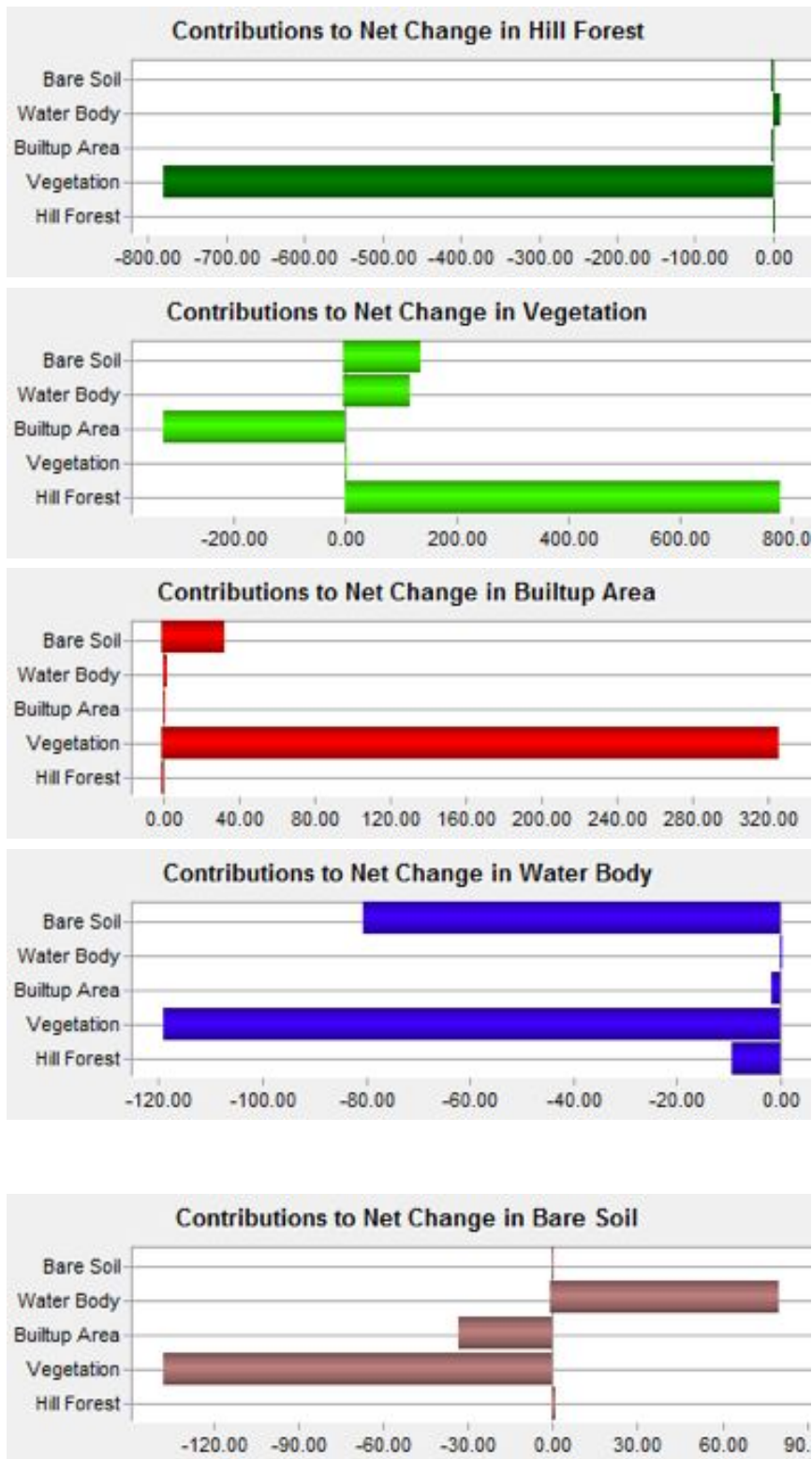




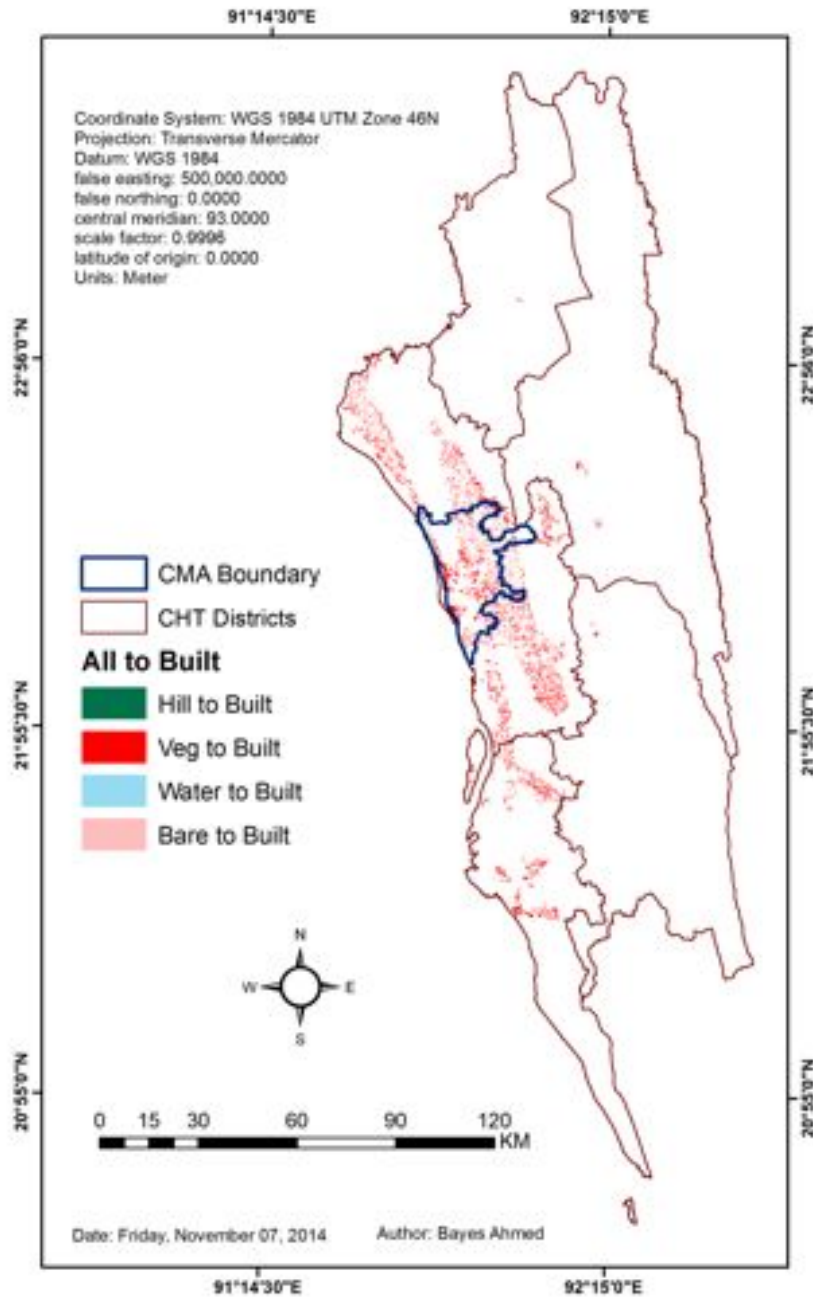


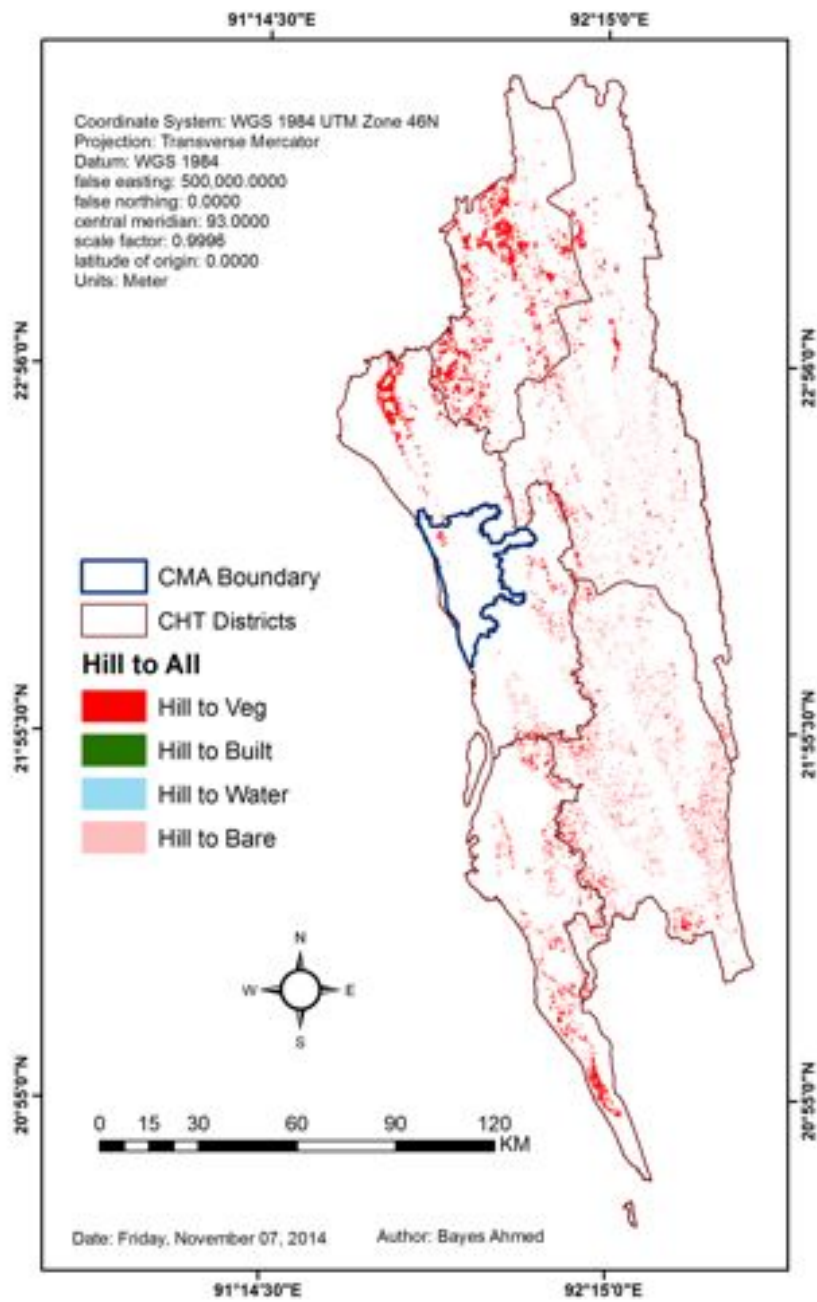


Appendix-II: Contributions to Different Land Cover Types in CHT (1990-2010)

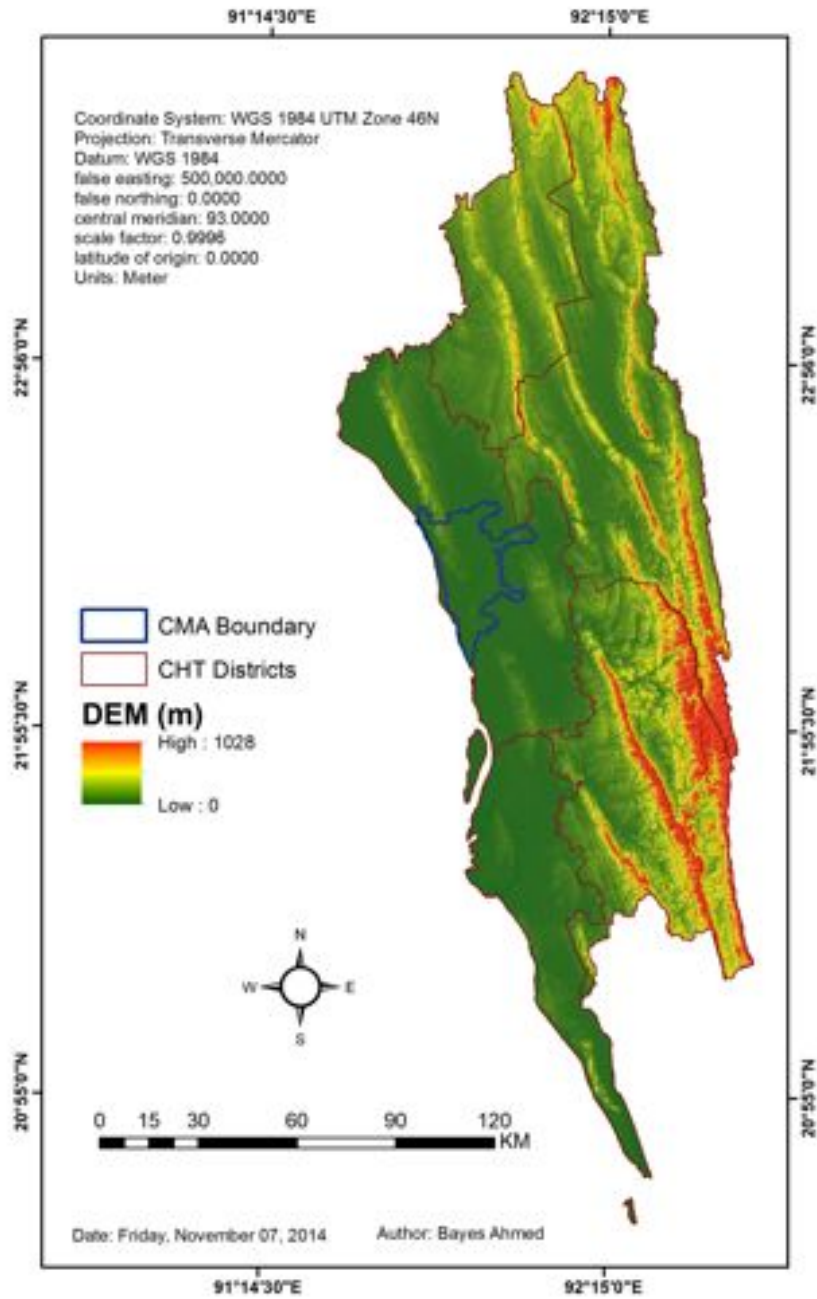


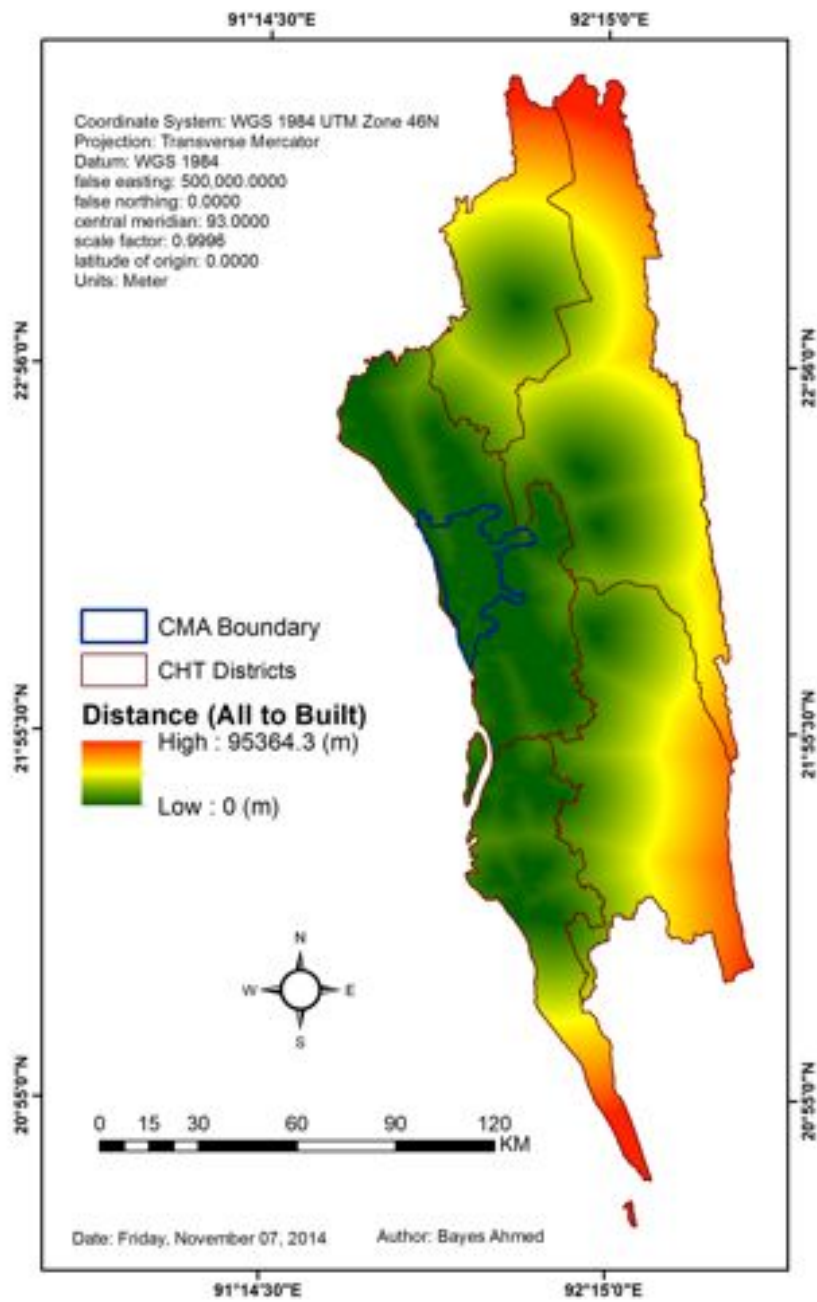
Appendix-III: Criteria Images for Land Cover Modelling

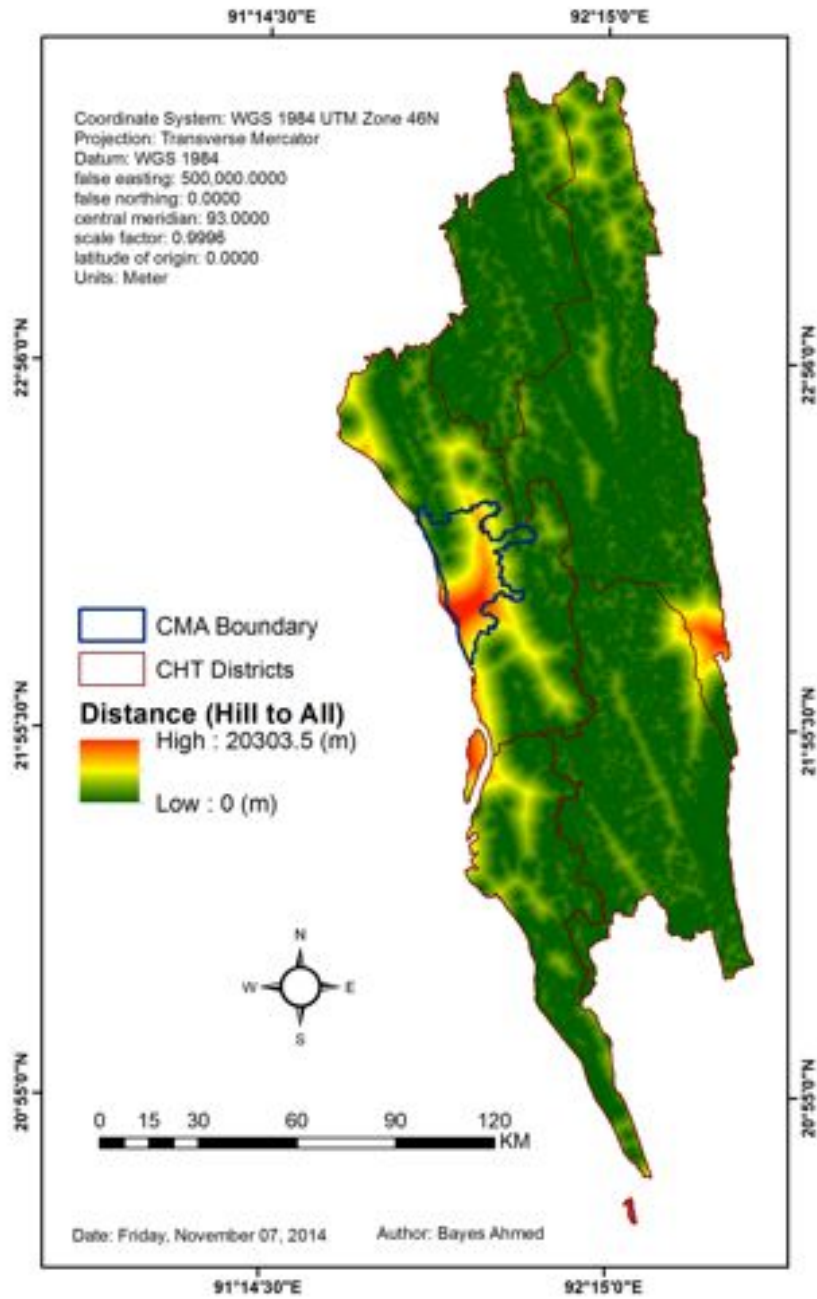


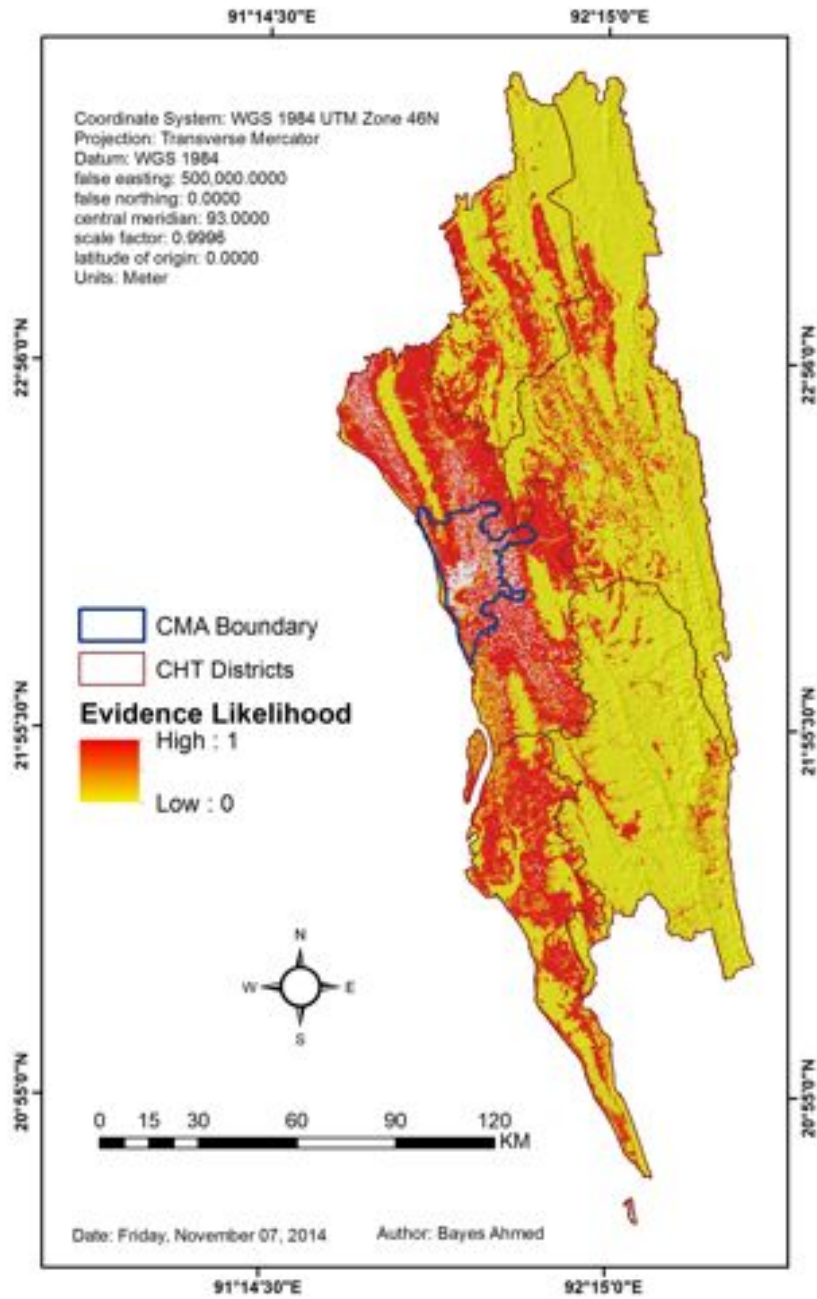


Appendix-IV: DEM and the Driving Variables for MLP_Markov Modelling









Appendix-V: Transition Potential Maps for Land Cover Modelling

